**MethodsCore**

**FirstLevel Script Documentation**

**09/05/2012**

**11/19/2013 - Updated**

Introduction

The FirstLevel script lets you specify a first level design. The script is as flexible as the SPM GUI, but will save you a lot of time. The FirstLevel Script assumes you’ve already created a MasterData file that describes the trials of your experiment. Once that is done, there are four major sections to fill out:

1. Image and subject information
2. MasterData file and condition information
3. User-specified regressor information
4. Contrast information

Additionally, there is an Advanced Section to the script, which only some options will be covered. There is a brief description of the advanced features in the script itself. If you want to use any of these features to the fullest extent, we recommend you contact [MethodsCoreHelp@umich.edu](mailto:MethodsCoreHelp@umich.edu) for information. Filling in these four sections in the script is mostly self-explanatory, and there is extensive help in the script itself. Below we’ll walk you through the steps that require more explanation.

Visual Tutorials

There are a series of video tutorials available online. It may be useful to view these while following along in this document. As of 11/19/13, parts of these videos are no longer relevant.

1 Introduction <http://youtu.be/FHWJgJFLMU8?hd=1>

2 MasterData File <http://youtu.be/6OoV_SPkzSU?hd=1>

3 Subjects & Runs <http://youtu.be/epfNW4u-z9c?hd=1>

4 Conditions, Onsets, & Durations <http://youtu.be/xsLuvXqwZy4?hd=1>

5 Parametric Regressors <http://youtu.be/SbXZ0CpxerE?hd=1>

6 User-Specified Regressors <http://youtu.be/D5kw4RJY1GU?hd=1>

7 Contrasts <http://youtu.be/EPSB2FObnyY?hd=1>

8 Conclusion <http://youtu.be/CrRaQELaKNw?hd=1>

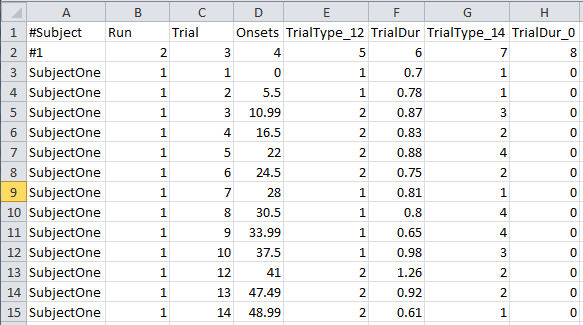
Getting Started

Copy the FirstLevel\_template script from your Methods Core directory into your local experiment directory. This is the file you’ll be editing.

Part 1: MasterData Files

The master data file is a CSV file that contains all the information about your experiment. All the information normally specified in the SPM GUI subject-by-subject and condition-by-condition is now specified in a single file. This makes analysis easier, more flexible, and more transparent. Its syntax has been revamped, so previous master data files are incompatible without minor modifications. The master data file can still be a trial-by-trial exhaustive list of every trial in an experiment and specifying the master data file in this manner will be discussed first. Additionally, the master data file can hold pointers to subject specific data files which can hold pointers to run specific files. Creating a master data file with pointers will be discussed second.

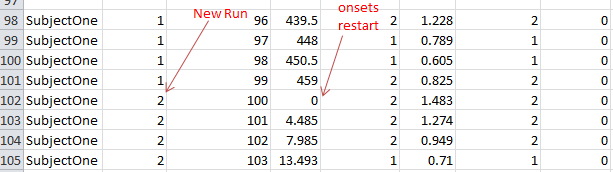
**Trial-by-trial data file:**



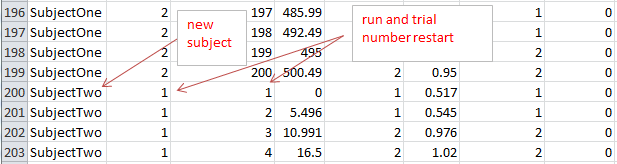
**Figure 1**

Look at Figure 1. I have two header rows here: row 1 identifies the type of column and row 2 provides the column number (you’ll see why this is useful a bit later). Notice the first column entries in the headers begin with the pound sign ‘#’. When placed at the beginning of a line, the ‘#’ character marks rows to ignore when reading data files (previously this was done by the MasterDataSkipRows variable in the first level template file which has been removed). The remaining content of the master data file is a trial-by-trial exhaustive list of every trial in the experiment. In the above spreadsheet, I have listed for each trial: the Subject (col 1), Run (col 2), Trial Number (col 3). I also have information for each trial about its Onset, Condition, and Duration but I’ll get to that a bit later.

The MasterData file can be one long list. When I get to the end of a run, I start the next run and continue listing trials. The onsets restart because SPM wants to know the onsets relative to the start of a run.



If you start a new subject, then you list the new subject number, and restart run number and trial number (and onsets).



Strings can now be present in the master data file except in the run, trial, onset, and duration columns. If they are present in these columns, an error will occur when reading the master data file. Strings are also used to identify subjects now. Of course you can have a string of numbers to identify subjects.

*Runs, Conditions, onsets, and durations*

For each trial we first assign it to a run. This trial should occur in that run during the experiment. In the master data file, the run numbers should be chosen such that they correspond with the runs to include vector in the SubjDir variable declared in the template file. The SubjDir variable will be discussed in more detail later.

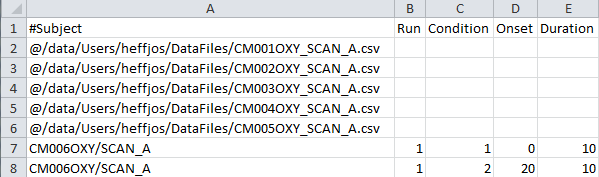
For each trial, we want to be able to assign to it a ***Condition***, an ***Onset***, and a ***Duration***. Recall that for an SPM analysis this information is required for every single trial. Assigning the ***Onset*** and ***Duration*** are easy – they are simply listed in Figure 1 in columns 4 and 6.

The ***Condition*** assignment is done with a numerical method. First, I assign each condition a number starting with 1. So let’s say I have two conditions, Congruent and Incongruent. Then I would assign Congruent as 1 and Incongruent as 2. Next, for each trial, I assign it the appropriate condition using these numbers. For example, in Figure 1, column 5, I assigned every trial a 1 or a 2 based on whether it is Congruent or Incongruent. A bit later, we will ‘point to’ the relevant columns in the MasterData file as we set up our FirstLevel script.

**Pointer data file:**

With the new data file syntax, pointers to other data files can be included in another data file, which permit the usage of multiple data files for a first level analysis. Users can now store run information at the subject and run level (Figure 2). The ‘@’ symbol indicates a line in a data file points to another data file (figure 3). Immediately after the ‘@’ symbol comes the full file path to the other data file. All data files need to have the subject, run, condition, onset, and duration information in the same columns; otherwise, the wrong information will probably be used to generate the first level results. Data files can store both data and pointers if the user desires to do such a thing.

**Figure 2.** Schematic outlining pointer data file method.

****

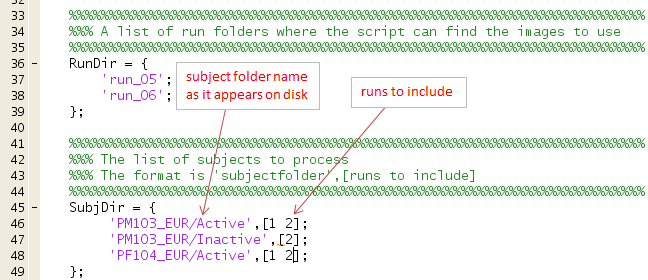
**Figure 3.** An example data file using pointer syntax (rows 2-6). Normal data is also included (rows 7-8), so this is a mixed data file.

Part 2: Setting up the FirstLevel\_template Script

You specify a lot of experiment-related information in the FirstLevel script. Most of this should be self-explanatory and there is help in the script itself.

One important issue is setting up paths. You set up the paths to a lot of things (path to functional images, output path, MasterData file, regressor file if you have one, etc…). Setting up paths is done through a standardized ‘Path Template’ system. See the **Path Template Documentation** in the root directory of the Methods Core repository for help on this. We’ll be using this method for setting up paths in pretty much all Methods Core software.

**Setting up Subjects and Runs**

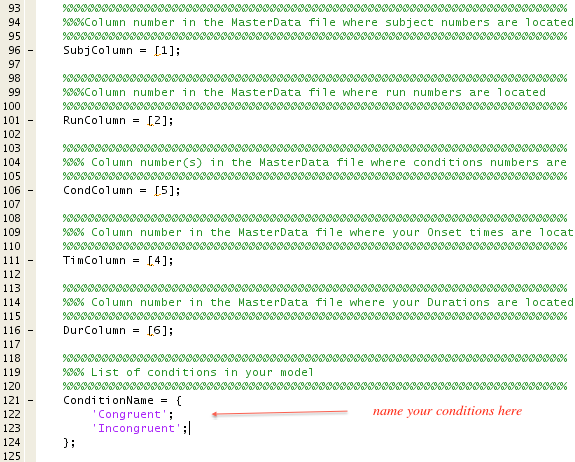


RunDir lists the run names as they appear on the disk. SubjDir lists the subjects to include in the analysis. It has two columns that are explained above. Notice that ‘runs to include’ (i.e., column 2 of SubjDir) indexes the rows of RunDir. So in the above example, we are omitting run\_05 for subject ‘PM103\_EUR/Inactive’.

**Setting up conditions, onsets, durations, etc..**

In the FirstLevel\_template Script, you ‘point to’ the relevant columns the MasterData file we set up earlier (see Figure 2 below). The script will then ‘find’ the information from the MasterData file. It will go to just the subjects, runs you tell it to, and will ignore the rest. Notice you also name your conditions, and you need to make sure the order of names corresponds to the order of numbers you used in the MasterData File.

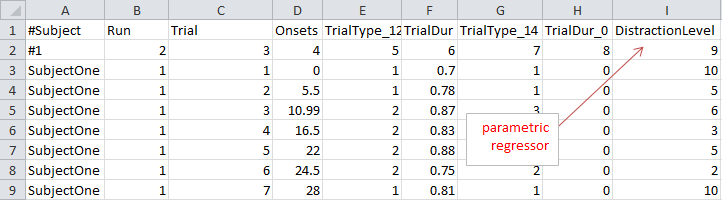
But let’s say I change my mind and want to model my data with four conditions: Congruent, Incongruent, Non-response, and Errors. Then I create a new condition column in the MasterData file with the numbers 1 through 4 that represent these conditions. This is column 7 in Figure 1. Then in the FirstLevel script (see Figure 4 below), I replace The CondColumn with 7, and list the new condition names in the ConditionName variable.



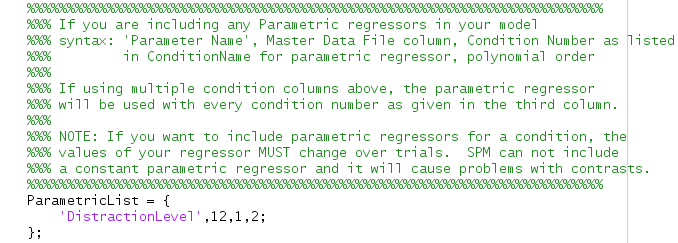
**Figure 4**

Parametric Regressors

A parametric regressor is a regressor that modulates activity on a trial-by-trial basis. Adding a parametric regressor is easy. Just add a new column in the MasterData file that specifies the value the parametric regressor takes for each trial.

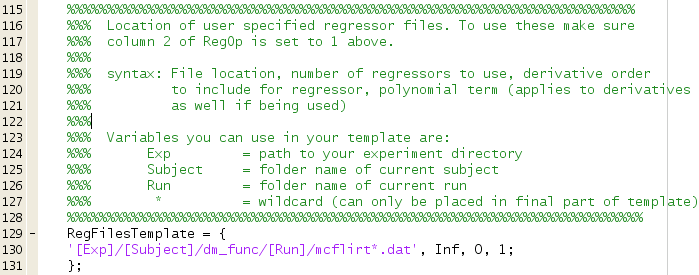


Then you list the name of the parametric regressor, its associated column in the FirstLevel script, the corresponding condition number, and the polynomial order (typically this will be 1).



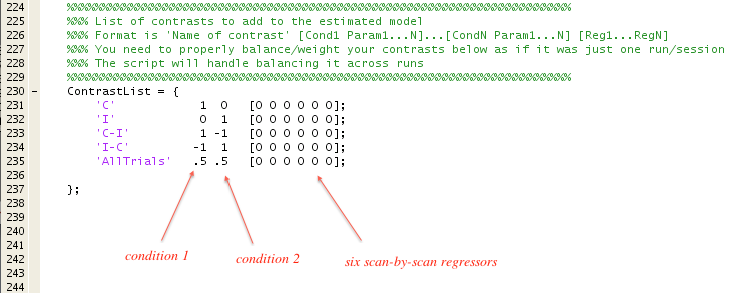
User-Specified Regressors

A user-specified regressor (the terminology is from SPM) is a regressor that modulates activity on a scan-by-scan basis. The most common scan-by-scan regressors are your motion regressors. So the FirstLevel script lets you specify scan-by-scan regressors by specifying a variable called RegFileTemplates. Each line of RegFileTemplates refers to a single regressor file, followed by a number indicating the number of columns you’d like to use from that file (use Inf to automatically use all the columns), followed by another number which indicates the derivative order to include for the regressor file (0 = no derivatives, 1 = first derivative, and so on), and final another number used to include the polynomial order for the file (1 = none, 2 = quadratic, 3 = cubic, and so on). The polynomial order is also applied to any derivatives calculated.

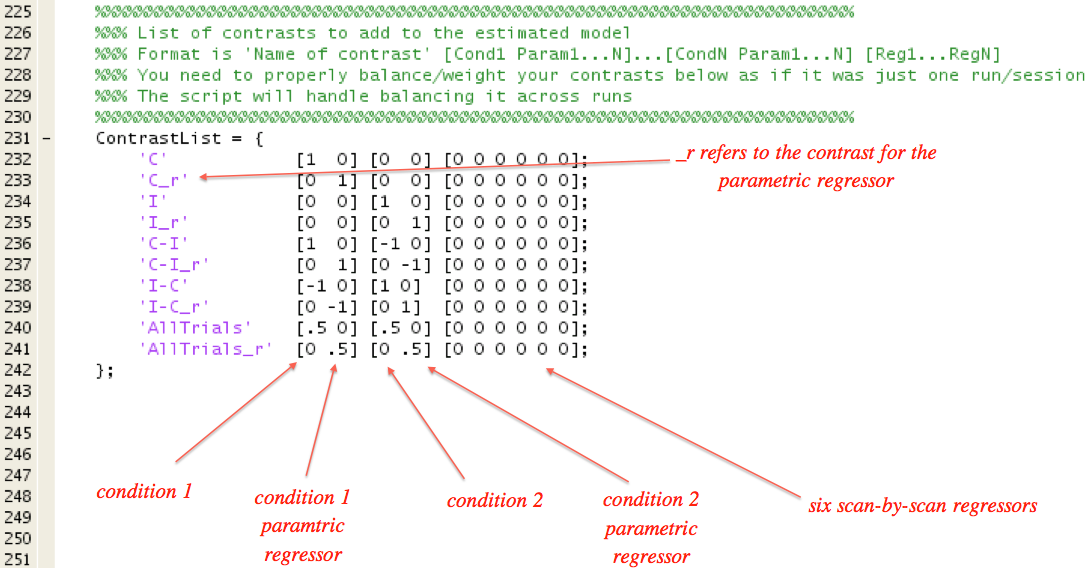


Contrasts

Let’s say you had an experiment with two conditions, Congruent and Incongruent, and 6 user-specified regressors. Here is what the contrast might look like:

****

Let’s say you now have 2 conditions and a single parametric regressor, and 6 user-specified regressors. Here is what the contrast might look like.



Brackets (i.e., [ and ]) that are placed around the condition/parametric regressor pair and also around the six scan-by-scan regressors are not for cosmetic purposes. The script uses the brackets to parse your ContrastList so do not omit the brackets!

Some other rules to keep in mind are (these are for most standard analyses, and there are exceptions):

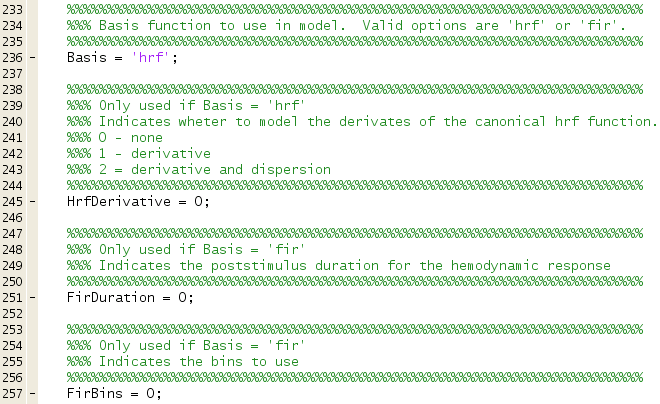
* Contrasts against an implicit baseline (e.g., ‘C’) – positive numbers sum to 1, no negative numbers
* Contrasts of two conditions against each other (e.g., ‘I-C’) - positive numbers sum to 1, negative numbers sum to 1

If you follow these rules, your estimated betas will correspond to percent signal change, which is a good thing. The script helps you follow these rules by balancing your numbers across runs. Here is how it works. You specify the contrasts as if there were only one run and you do not include any trailing zeros for the SPM-created run-related regressor. The script will automatically expand/rescale the numbers for the number of runs in your experiment. For example, if there are three runs, all the numbers in the contrasts above would be multiplied by 1/3, and additional trailing zeros for the SPM-created run-related regressors (one per run) would be automatically be created.

To repeat, there is also an Advanced Section, with a LOT of other features. For example, you can have run specific contrasts. You reset SPM defaults. You can drop conditions that appear just a few times. And many others. If you want to use any of these features, we recommend you contact [MethodsCoreHelp@umich.edu](mailto:MethodsCoreHelp@umich.edu) for information.

**Advanced: Basis functions**

In this context, basis functions are the assumed brain response caused by an external stimulus (i.e. a trial in an experiment). The advanced section gives the user flexibility choosing which basis functions to use in a first level analysis (Figure 5). With the ‘Basis’ variable, you can choose between using the canonical HRF or finite impulse response (FIR) basis functions. When using the canonical HRF, you can additionally model the time and dispersion derivatives using the ‘HrfDerivative’ variable. If you do model the derivatives, this will affect how you setup your contrasts. Example contrasts will be given later.

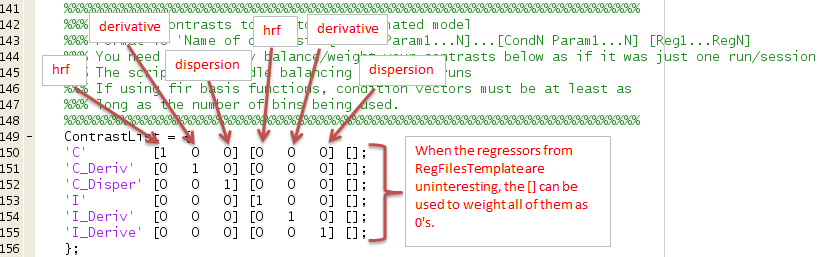


**Figure 5.** The variables used for setting basis functions in the advanced section.

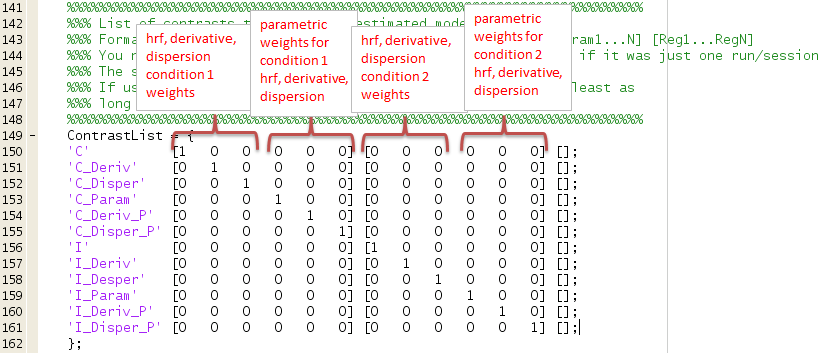
To use an FIR model in the first level, set ‘Basis’ to ‘fir’. The FIR model is commonly used to estimate a condition time course in an experiment. The post-stimulus duration is controlled by ‘FirDuration’ and the number of basis functions is controlled by the ‘FirBins’ variable. If we wanted to estimate a 30 second post-stimulus time course with 15 FIR basis functions, we set FirDuation=30 and FirBins=15.

**Advanced: Contrasts**

Let’s say we have 2 conditions and are using the ‘hrf’ basis with the derivative and dispersion regressors.



Let’s say we have 2 conditions, are using the ‘hrf’ basis with the derivative and dispersion regressors, and are using parametric regressors for both conditions. The contrasts will be similar to above, except three additional columns are added for both condition vectors for the parametric regressors.



Now let’s say we have 2 conditions, are using the ‘fir’ basis function with an order equal to 10. We will have 10 weights for each condition.

